

Figure 1.

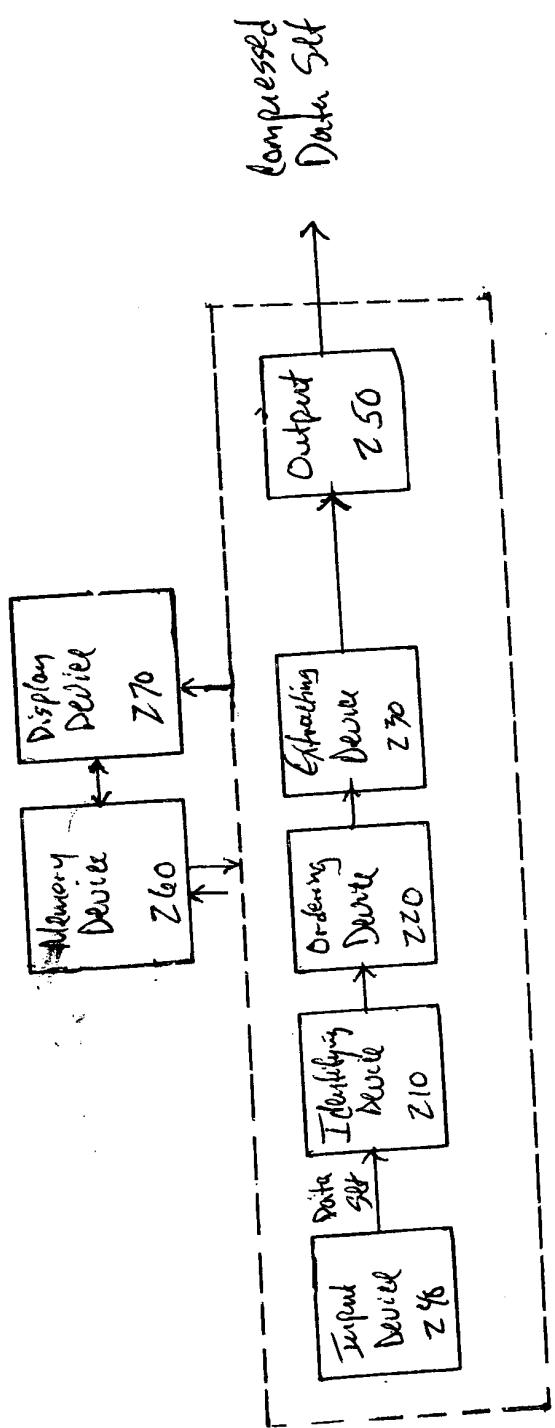


Figure 2.

	1	2	3	4	5
$m_1$	a	b	c	.	d
$m_2$	a	b	e		
$m_3$	a	d	d	.	d
$m_4$	a	d	.	.	e
$m_5$	a	b	a	.	d

Figure 3

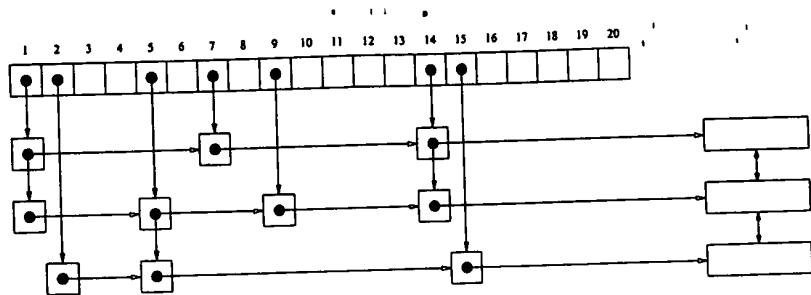


Figure 1: Threading motifs and their occurrences.

Figure 4

Table 1: Lossy compression of gray-scale images (1 pixel = 1 byte).

file	file len	GZip len [%compr]	<i>Codec<sub>2</sub></i> [%compr]	<i>Codec<sub>1</sub></i> [%compr]	%Diff gzip	%loss	‘.’/ char
bridge	66336	61657 <sub>[7.05]</sub>	60987 <sub>[8.06]</sub>	57655 <sub>[13.08]</sub>	6.49	0.42	1/4
			60987 <sub>[8.06]</sub>	50656 <sub>[23.63]</sub>	17.84	14.29	1/3
camera	66336	48750 <sub>[26.51]</sub>	47842 <sub>[27.88]</sub>	46192 <sub>[30.36]</sub>	5.25	0.74	1/6
			48044 <sub>[27.57]</sub>	45882 <sub>[30.83]</sub>	5.88	2.17	1/5
lena	262944	234543 <sub>[12.10]</sub>	226844 <sub>[13.73]</sub>	210786 <sub>[19.83]</sub>	11.60	9.09	1/4
			186359 <sub>[29.13]</sub>	175126 <sub>[33.39]</sub>	25.33	20.00	1/3
peppers	262944	232334 <sub>[11.64]</sub>	218175 <sub>[17.03]</sub>	199605 <sub>[23.85]</sub>	14.09	6.25	1/4
			180783 <sub>[31.25]</sub>	173561 <sub>[33.99]</sub>	25.30	20.00	1/3

Figure 5.

Table 2: Lossy compression of binary images.

file	file len	GZip len [%compr]	Codec <sub>2</sub> [%compr]	Codec <sub>1</sub> [%compr]	%Diff GZip	%loss	':/' char
ccitt7	513229	109612 <sub>[78.64]</sub>	98076 <sub>[80.89]</sub>	91399 <sub>[82.19]</sub>	16.62	16.67	1/5
			93055 <sub>[81.87]</sub>	90873 <sub>[82.29]</sub>	17.10	16.67	1/4
test4	279213	58736 <sub>[78.96]</sub>	92658 <sub>[81.95]</sub>	85391 <sub>[83.36]</sub>	22.10	25.00	1/3
			57995 <sub>[79.23]</sub>	54651 <sub>[80.42]</sub>	6.95	0.91	1/4
			57714 <sub>[79.32]</sub>	54402 <sub>[80.51]</sub>	7.38	1.27	1/3

Figure 6

Table 3: Lossy compression of music (1 sample = 1 byte).

file	file len	GZip len [%compr]	Codec <sub>2</sub> [%compr]	Codec <sub>1</sub> [%compr]	%Diff GZip	%loss	':/' char
crowd	128900	103834 <sub>[19.44]</sub>	92283 <sub>[28.41]</sub>	86340 <sub>[33.01]</sub>	16.85	16.67	1/3
	196834	171846 <sub>[12.96]</sub>	148880 <sub>[24.36]</sub>	139308 <sub>[29.22]</sub>	18.93	9.09	1/4
			114709 <sub>[41.72]</sub>	111058 <sub>[43.57]</sub>	35.37	25.00	1/3

Figure 7

Table 4: Lossy vs. lossless performance.

file	file dim	GZip [%compr]	Codec <sub>1</sub> [%compr]	%loss	':/' char	Lossless [%compr]	%Diff GZip
bridge	66336	61657 <sub>[7.05]</sub>	50656 <sub>[23.63]</sub>	14.29	1/3	59344 <sub>[10.54]</sub>	3.75
	66336	48750 <sub>[26.51]</sub>	43096 <sub>[35.03]</sub>	9.09	1/4	45756 <sub>[31.02]</sub>	6.14
	262944	234543 <sub>[12.10]</sub>	175126 <sub>[33.39]</sub>	20.00	1/3	199635 <sub>[24.07]</sub>	14.88
	262944	232334 <sub>[11.64]</sub>	199605 <sub>[23.85]</sub>	6.25	1/4	211497 <sub>[19.56]</sub>	8.97
camera			173561 <sub>[33.99]</sub>	20.00	1/3	195744 <sub>[25.55]</sub>	15.75
	66336	48750 <sub>[26.51]</sub>	43096 <sub>[35.03]</sub>	9.09	1/4	45756 <sub>[31.02]</sub>	6.14
	262944	234543 <sub>[12.10]</sub>	175126 <sub>[33.39]</sub>	20.00	1/3	199635 <sub>[24.07]</sub>	14.88
	262944	232334 <sub>[11.64]</sub>	199605 <sub>[23.85]</sub>	6.25	1/4	211497 <sub>[19.56]</sub>	8.97
lena	66336	61657 <sub>[7.05]</sub>	50656 <sub>[23.63]</sub>	14.29	1/3	59344 <sub>[10.54]</sub>	3.75
	66336	48750 <sub>[26.51]</sub>	43096 <sub>[35.03]</sub>	9.09	1/4	45756 <sub>[31.02]</sub>	6.14
	262944	234543 <sub>[12.10]</sub>	175126 <sub>[33.39]</sub>	20.00	1/3	199635 <sub>[24.07]</sub>	14.88
	262944	232334 <sub>[11.64]</sub>	199605 <sub>[23.85]</sub>	6.25	1/4	211497 <sub>[19.56]</sub>	8.97
peppers	66336	61657 <sub>[7.05]</sub>	50656 <sub>[23.63]</sub>	14.29	1/3	59344 <sub>[10.54]</sub>	3.75
	66336	48750 <sub>[26.51]</sub>	43096 <sub>[35.03]</sub>	9.09	1/4	45756 <sub>[31.02]</sub>	6.14
	262944	234543 <sub>[12.10]</sub>	175126 <sub>[33.39]</sub>	20.00	1/3	199635 <sub>[24.07]</sub>	14.88
	262944	232334 <sub>[11.64]</sub>	199605 <sub>[23.85]</sub>	6.25	1/4	211497 <sub>[19.56]</sub>	8.97
ccitt7	513229	109612 <sub>[78.64]</sub>	90873 <sub>[82.29]</sub>	16.67	1/4	97757 <sub>[80.09]</sub>	10.82
	513229	109612 <sub>[78.64]</sub>	85391 <sub>[83.36]</sub>	25.00	1/3	89305 <sub>[82.59]</sub>	18.53
	279213	58736 <sub>[78.96]</sub>	54402 <sub>[80.51]</sub>	1.27	1/3	54875 <sub>[80.34]</sub>	6.57
	279213	58736 <sub>[78.96]</sub>	54402 <sub>[80.51]</sub>	1.27	1/3	54875 <sub>[80.34]</sub>	6.57
test4	128900	103834 <sub>[19.44]</sub>	86340 <sub>[33.01]</sub>	16.67	1/3	96903 <sub>[24.82]</sub>	6.68
	196834	171846 <sub>[12.96]</sub>	139308 <sub>[29.22]</sub>	9.09	1/4	159206 <sub>[19.11]</sub>	7.36
	196834	171846 <sub>[12.96]</sub>	139308 <sub>[29.22]</sub>	9.09	1/4	159206 <sub>[19.11]</sub>	7.36
			111058 <sub>[43.57]</sub>	25.00	1/3	151584 <sub>[22.98]</sub>	11.97

Figure 8

Table 5: Lossless compression of Calgary Corpus.

file	file len	GZip [%compr]	Codec <sub>1</sub> [%compr]	%loss	':/char	Lossless [%compr]	%Diff GZip
bib	111261	35063 <sub>[68.49]</sub>	36325 <sub>[67.35]</sub>	3,70	1/3	37491 <sub>[66.30]</sub>	6.92
book1	768771	313376 <sub>[60.01]</sub>	245856 <sub>[68.01]</sub>	12.50	1/3	277180 <sub>[63.95]</sub>	11.55
book2	610856	206687 <sub>[66.16]</sub>	197199 <sub>[67.72]</sub>	4.35	1/4	202713 <sub>[66.81]</sub>	1.92
geo	102400	68493 <sub>[33.11]</sub>	40027 <sub>[60.91]</sub>	16.67	1/4	63662 <sub>[37.83]</sub>	7.05
news	377109	144840 <sub>[61.59]</sub>	144541 <sub>[61.67]</sub>	0.42	1/3	144644 <sub>[61.64]</sub>	0.14
obj1	21504	10323 <sub>[51.99]</sub>	8386 <sub>[61.00]</sub>	16.67	2/5	9221 <sub>[57.12]</sub>	10.68
obj2	246814	81631 <sub>[66.93]</sub>	71123 <sub>[71.18]</sub>	20.00	1/2	83035 <sub>[66.36]</sub>	-1.72
paper1	53161	18577 <sub>[65.06]</sub>	19924 <sub>[62.52]</sub>	1.75	1/3	20174 <sub>[62.05]</sub>	-8.60
paper2	82199	29753 <sub>[63.80]</sub>	29920 <sub>[63.60]</sub>	0.76	1/2	30219 <sub>[63.24]</sub>	-1.57
pic	513216	56422 <sub>[89.01]</sub>	52229 <sub>[89.82]</sub>	0.56	1/3	52401 <sub>[89.79]</sub>	7.13
progC	39611	13275 <sub>[66.49]</sub>	13840 <sub>[65.06]</sub>	1.32	1/2	14140 <sub>[64.30]</sub>	-6.52
progl	71646	16273 <sub>[77.29]</sub>	17249 <sub>[75.92]</sub>	0.58	1/3	17355 <sub>[75.78]</sub>	-6.65
progp	49379	11246 <sub>[77.23]</sub>	12285 <sub>[75.12]</sub>	0.64	1/3	12427 <sub>[74.83]</sub>	-10.50

Figure 9

Table 6: Lossless compression of sequences from DNA yeast families.

file	file len	GZip [%compr]	Codec <sub>1</sub> [%compr]	%loss	':/char	Lossless [%compr]	%Diff GZip
Spor EarlyII	25008	8008 <sub>[67.98]</sub>	6990 <sub>[72.05]</sub>	0.45	1/3	7052 <sub>[71.80]</sub>	11.94
Spor EarlyI	31039	9862 <sub>[68.23]</sub>	8845 <sub>[71.50]</sub>	0.36	1/3	8914 <sub>[71.28]</sub>	9.61
Helden CGN	32871	10379 <sub>[68.43]</sub>	8582 <sub>[73.89]</sub>	1.33	1/3	8828 <sub>[73.14]</sub>	14.94
Spor Middle	54325	16395 <sub>[69.82]</sub>	14839 <sub>[72.68]</sub>	0.36	1/4	14924 <sub>[72.53]</sub>	8.97
Helden All	112507	33829 <sub>[69.93]</sub>	29471 <sub>[73.81]</sub>	1.56	1/4	29862 <sub>[73.46]</sub>	11.73
Spor All	222453	68136 <sub>[69.37]</sub>	56323 <sub>[74.68]</sub>	1.61	1/3	57155 <sub>[74.31]</sub>	16.12
All Up 400k	399615	115023 <sub>[71.22]</sub>	93336 <sub>[76.64]</sub>	14.29	1/3	106909 <sub>[73.25]</sub>	7.05

Figure 10

Table 7: Synopsis of compression rates for sequences in the yeast DNA by various lossless methods. The figure in parenthesis is the percentage gain of  $Codec_1$  versus other methods.

File	File Len	Huffman Pack [%diff]	LZ-78 Compress [%diff]	LZ-77 GZip [%diff]	BWT BZip [%diff]	$Codec_1$ Lossless
Spor EarlyII	25008	7996 <sub>[13.4]</sub>	7875 <sub>[11.7]</sub>	8008 <sub>[13.6]</sub>	7300 <sub>[3.5]</sub>	7052
Spor EarlyI	31039	9937 <sub>[11.5]</sub>	9646 <sub>[8.2]</sub>	9862 <sub>[10.6]</sub>	9045 <sub>[1.5]</sub>	8914
Helden CGN	32871	10590 <sub>[20.0]</sub>	10223 <sub>[15.8]</sub>	10379 <sub>[17.6]</sub>	9530 <sub>[8.0]</sub>	8828
Spor Middle	54325	17295 <sub>[15.9]</sub>	16395 <sub>[9.9]</sub>	16395 <sub>[9.9]</sub>	15490 <sub>[3.8]</sub>	14924
Helden All	112507	36172 <sub>[21.1]</sub>	33440 <sub>[12.0]</sub>	33829 <sub>[13.3]</sub>	31793 <sub>[6.5]</sub>	29862
Spor All	222453	70755 <sub>[23.8]</sub>	63939 <sub>[11.9]</sub>	68136 <sub>[19.2]</sub>	61674 <sub>[7.9]</sub>	57155
All Up 400k	399615	121700 <sub>[13.8]</sub>	115029 <sub>[7.6]</sub>	115023 <sub>[7.6]</sub>	112363 <sub>[5.1]</sub>	106909

Figure 11

Table 8: Compression, fidelity and loss in reconstruction of grey scale images.

File	File len	GZip len [%compr]	$Codec_1$ [%compr]	Diff % GZip	%Loss	'.'/ car	%Loss per pixel over recon pix	%Loss per pixel over all pix
bridge	66336	61657 <sub>[7.05]</sub>	57655 <sub>[13.08]</sub> 60987 <sub>[8.06]</sub>	6.49 17.84	0.42 14.29	1/4 1/3	5.67 7.69	0.02 0.90
camera	66336	48750 <sub>[26.51]</sub>	47316 <sub>[28.67]</sub>	11.60	9.09	1/4	0.78	0.05
lena	262944	234543 <sub>[12.10]</sub>	210786 <sub>[19.83]</sub> 186359 <sub>[29.13]</sub>	10.13 25.33	4.17 20.00	1/4 1/3	7.26 5.11	0.27 0.81
peppers	262944	232334 <sub>[11.64]</sub>	199605 <sub>[23.85]</sub> 180783 <sub>[31.25]</sub>	14.09 25.30	6.25 20.00	1/4 1/3	1.53 3.29	0.08 0.52

Figure 12

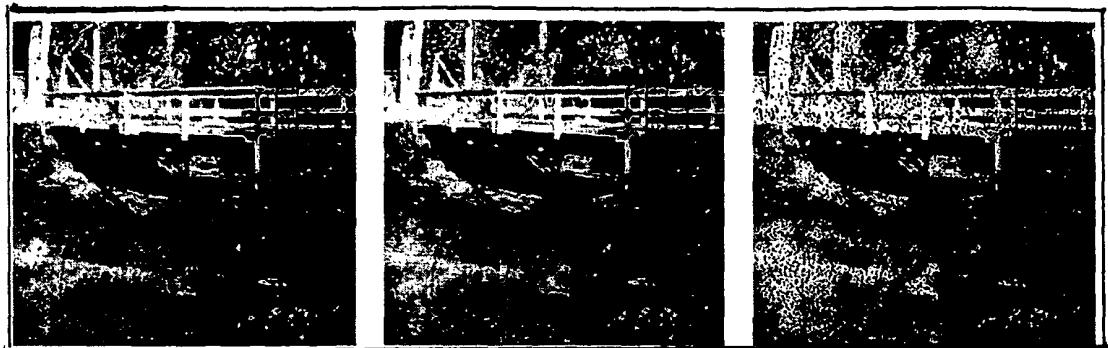


Figure  
13A



Figure  
13B



Figure  
13C

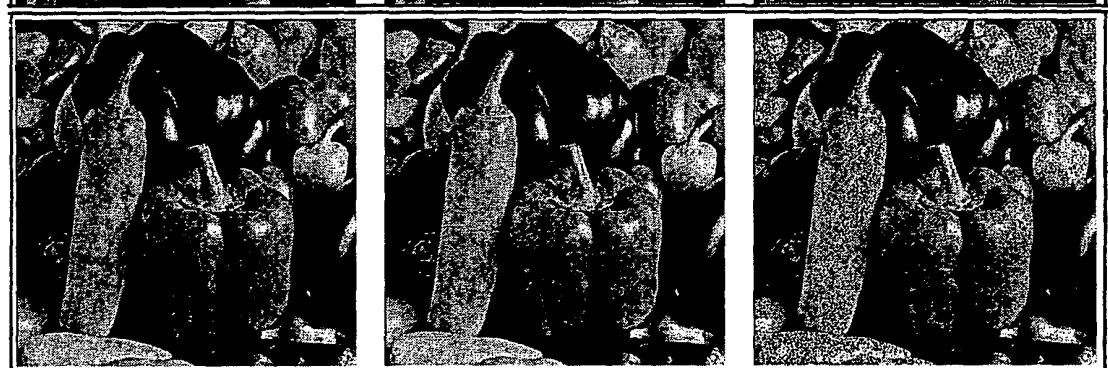


Figure  
13D

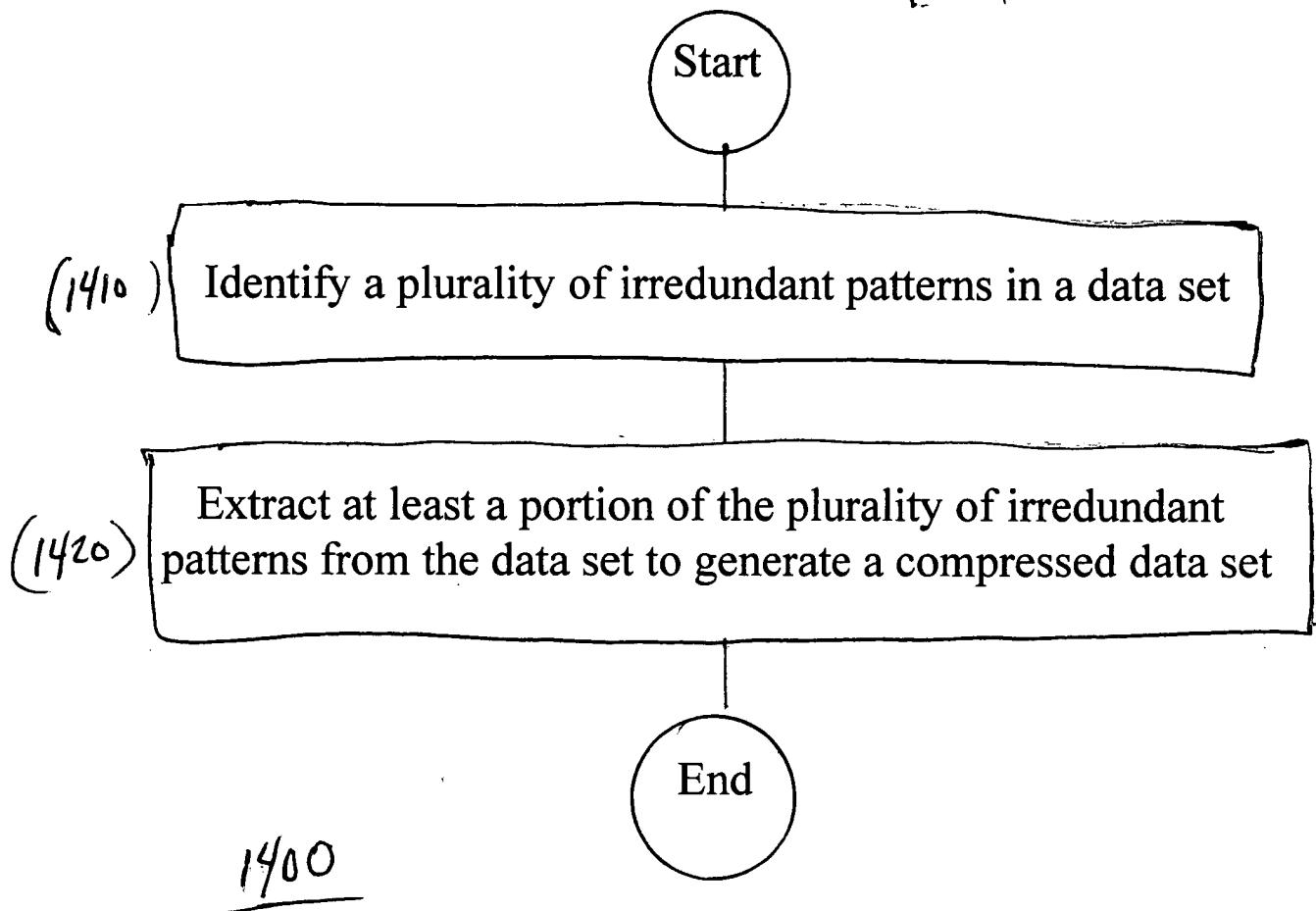


Figure 14

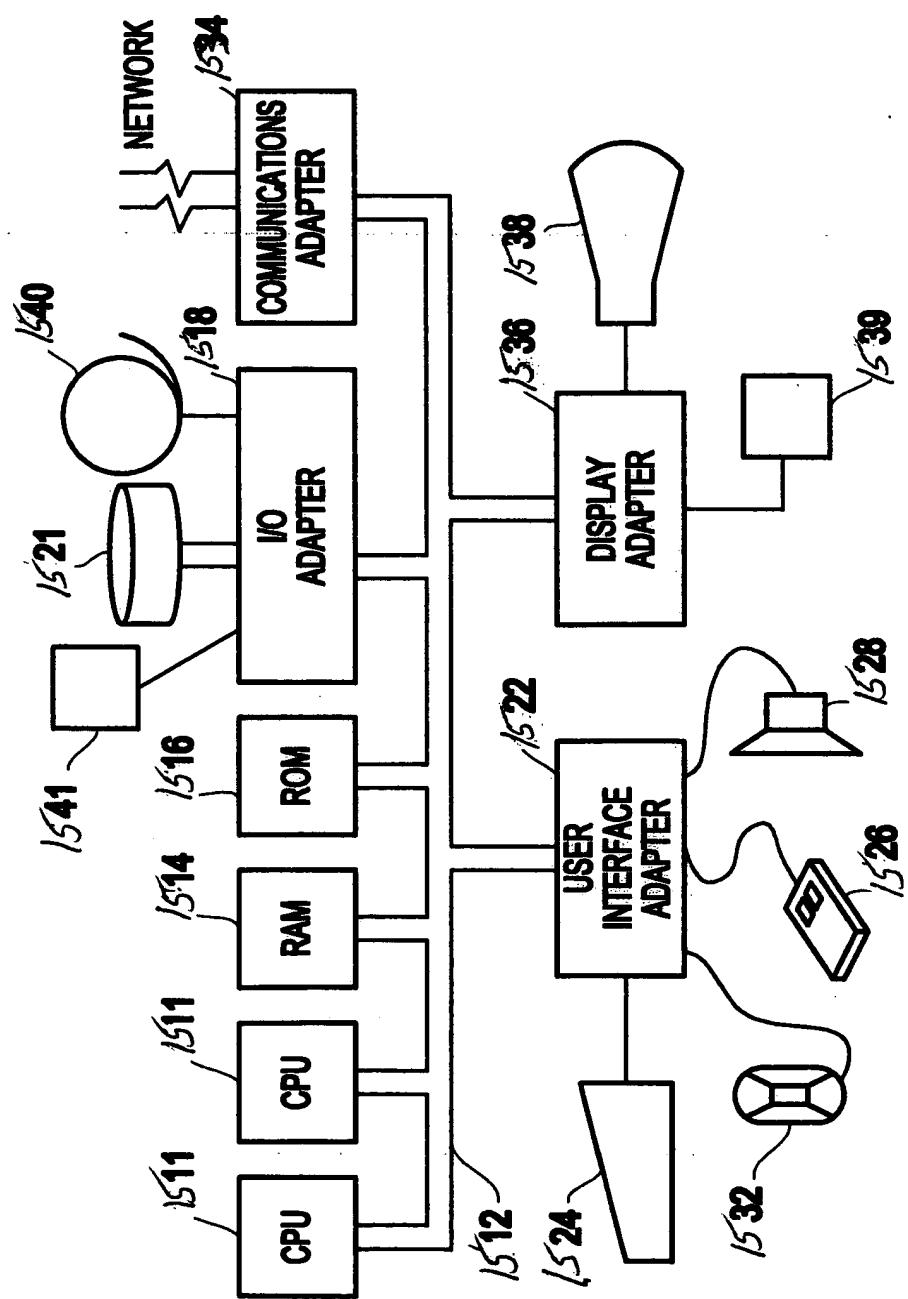


FIG. 15

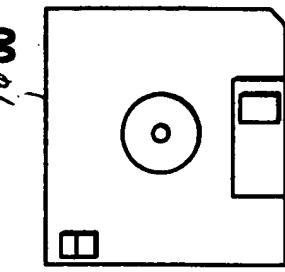


FIG. 16

**DESCRIPTION OF FARMER OAK - AN INCIDENT** When Farmer Oak smile, the corners of his mouth spread till they were within an unimportant distance of his ears, his eyes were reduced to chinks, and diverging wrinkles round them, extending upon ... countenance like the rays in a rudimentary sketch of the rising sun. His Christian name was Gabriel, and on working days he was a young man of sound judgment, easy motions, proper dress, and ... general good character. On Sundays, he was a man of misty views rather given to postponing, and hampered by his best clothes and umbrella : upon ... whole, one who felt himself to occupy morally that ... middle space of Laodicean neutrality which ... between the Communion people of the parish and the drunken section, - that ... he went to church, but yawned privately by the time the congregation reached the Nicene creed, - and thought of what there would be for dinner when he meant to be listening to the sermon.

**DESCRIPTION OF FARMER OAK - AN INCIDENT** When Farmer Oak smiled, the corners of his mouth spread till they were within an unimportant distance of his ears, his eyes were reduced to chinks, and diverging wrinkles appeared round them, extending upon his countenance like the rays in a rudimentary sketch of the rising sun. His Christian name was Gabriel, and on working days he was a young man of sound judgment, easy motions, proper dress, and general good character. On Sundays he was a man of misty views, rather given to postponing, and hampered by his best clothes and umbrella : upon the whole, one who felt himself to occupy morally that vast middle space of Laodicean neutrality which lay between the Communion people of the parish and the drunken section, - that is, he went to church, but yawned privately by the time the congregation reached the Nicene creed, - and thought of what there would be for dinner when he meant to be listening to the sermon.

Figure 17